



**Cultura**

Revista de História e Teoria das Ideias

**Vol. 26 | 2009**

**O Tempo das Revistas**

---

## Edição electrónica e estudo de revistas

o contributo do Seminário Livre de História das Ideias

*Electronic publishing of magazines: tendencies, limitations and the contribution of the SLHI*

**Pedro Lisboa**

---



### Edição electrónica

URL: <http://journals.openedition.org/cultura/536>

DOI: 10.4000/cultura.536

ISSN: 2183-2021

### Editora

Centro de História da Cultura

### Edição impressa

Data de publicação: 1 Junho 2009

Paginação: 259-268

ISSN: 0870-4546

### Referência eletrónica

Pedro Lisboa, « Edição electrónica e estudo de revistas », *Cultura* [Online], Vol. 26 | 2009, posto online no dia 16 setembro 2013, consultado a 21 abril 2019. URL : <http://journals.openedition.org/cultura/536> ; DOI : 10.4000/cultura.536

---

## **Edição electrónica e estudo de revistas: o contributo do Seminário Livre de História das Ideias**

*Pedro Lisboa\**

A edição electrónica de revistas vive actualmente um momento de experiências e avanços. A principal e mais palpável consequência deste facto é que a pesquisa electrónica permite, na esmagadora maioria dos casos, o manuseamento do texto em formato digital, permitindo assim também, através da pesquisa livre, a verificação automática das ocorrências do termo pesquisado. Como a sua designação indica, este tipo de busca não implica qualquer pressuposto teórico organizativo: a interacção entre o leitor e o texto não tem qualquer tipo de mediação. Esta atitude perante a informação é estruturada pelo princípio geral do acesso directo às fontes, que, sem ser redutível à questão da pesquisa, direcciona em grande parte a discussão e os desenvolvimentos na área.

A este respeito, pode dizer-se que existem hoje duas grandes correntes de edição electrónica de revistas: uma, que resume o seu serviço à pesquisa livre directa nos textos<sup>1</sup>; outra, que acrescenta a opção de pesquisa direccionada, com palavras-chave, assuntos, etc.<sup>2</sup>. O elemento comum às duas é, assim, a pesquisa livre, tornada onnipresente e indispensável no quotidiano da investigação científica (e mesmo no quotidiano, sem mais!) por motores de busca de Internet como o Google. A grande diferença objectiva entre o acesso à informação em papel ou em suporte electrónico encontra-se na utilização do computador como meio de integração da pesquisa e da leitura – separando o modelo geralmente seguido na indexação e classificação biblioteconómicas clássicas, em utilização nas bibliotecas físicas, do da publicação electrónica.

A principal desvantagem do uso exclusivo da pesquisa livre passa por, como refere Jeffrey Garrett, ficarmos dependentes de uma “contagem cega de palavras-chave

\* Seminário Livre de História das Ideias do Centro de História da Cultura da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa.

<sup>1</sup> Cf., por exemplo, a JSTOR (<http://www.jstor.org>) ou a Highwire Press da Stanford University (<http://highwire.stanford.edu>).

<sup>2</sup> Cf., por exemplo, a SciELO (<http://www.scielo.org>) ou o Project Euclid (<http://projecteuclid.org>).

colhidas de forma mecânica” – e, poderíamos acrescentar, automática. Perante este cenário, e ainda segundo o mesmo autor, a pesquisa por termos analíticos funciona como “o dique que erigimos contra a enchente de informação que a progressiva digitalização de todos os textos está a criar. Ou antes, com mais precisão, é a torneira que colocámos nesse dique”<sup>3</sup>.

Sem nunca excluir a via da pesquisa livre, proveitosa em algumas situações, o Seminário Livre de História das Ideias (SLHI) pretende oferecer um acesso direccionado à informação, partindo dos princípios de representação descritiva da biblioteconomia e inovando e adaptando-os, de forma a reunir o melhor de dois mundos. Assim, o presente artigo procurará descrever o trabalho do Seminário sobre a edição electrónica de revistas, bem como discutir algumas das vantagens, desvantagens e implicações de um modelo híbrido de edição e estudo face às tendências actuais e aos objectivos propostos.

As revistas trabalhadas pelo SLHI inserem-se, maioritariamente, num plano concebido para representar as grandes correntes ideológicas dos movimentos intelectuais portugueses do princípio do século XX. Assim, são privilegiados os títulos proeminentes do republicanismo e da Renascença Portuguesa (*A Águia*, *Vida Portuguesa*, *Pela Grei*, etc.), do Integralismo Lusitano (*Nação Portuguesa*, *Integralismo Lusitano*, *Alma Portuguesa*), do pensamento católico (*Estudos*, *Estudos Sociais*), do anarquismo (*Suplemento Literário e Ilustrado* de *A Batalha*, *Renovação*, *A Sementeira*, *Germinal*) e do pensamento feminino (*Alma Feminina*, *Educação Feminina*). A par destes, juntam-se também algumas publicações que, fugindo ao critério geral de inclusão, foram escolhidas pela sua relevância contextual ou temática (*O Tempo e o Modo*, *Revista de Educação Geral e Técnica*, *Revista da Índia*, *Índia Nova*). Desde logo, ficam delimitadas as fontes: revistas portuguesas emblemáticas de ideias e cultura, maioritariamente do começo do século passado. As particularidades deste *corpus* não são alheias à construção da metodologia, que se quer adaptada ao objecto de trabalho.

O processo de trabalho inicia-se com a recolha de dados biográficos gerais acerca da publicação. O investigador (ou investigadores) responsável reúne toda a informação pertinente sobre a vida editorial da revista, como o seu período e local de publicação, os seus colaboradores mais destacados e o seu corpo redactorial. Pretende-se que, antes de mais, seja traçada uma primeira história do percurso do periódico, mesmo

<sup>3</sup> Jeffrey Garrett, “KWIC and Dirty? Human Cognition and the Claims of Full-Text Searching”, in *Journal of Electronic Publishing*, Vol. 9, n.º 1, Ann Arbor, University of Michigan, Inverno de 2006.

sendo passível de posteriores correcções e acrescentos decorrentes do curso do trabalho, que permita formar alicerces factuais. A par deste esforço, são feitas leituras complementares, quer de outras fontes quer de estudos, conferindo ao investigador uma contextualização plena.

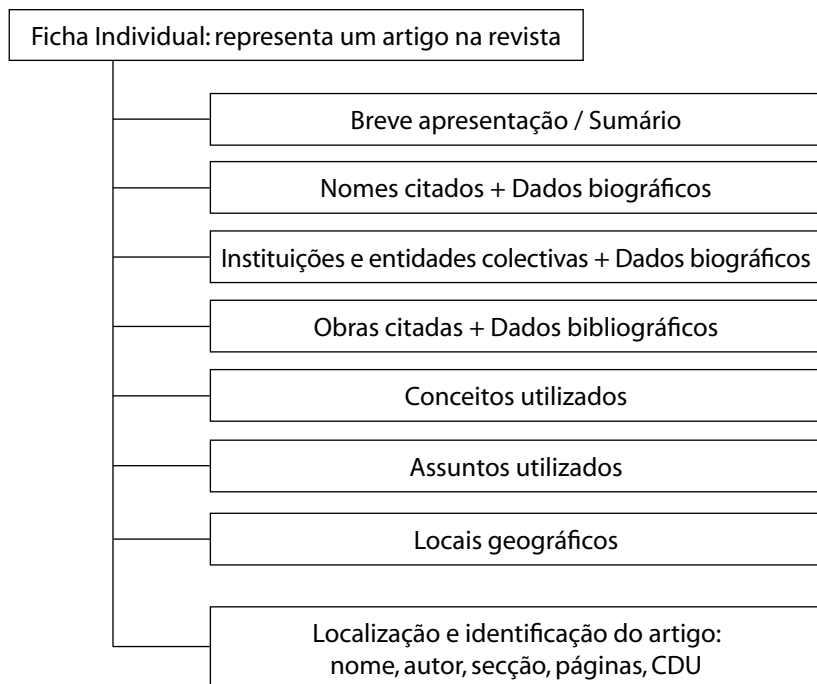
Segue-se uma leitura exaustiva da revista na sua totalidade. Este passo é de uma evidente relevância para a apreensão do significado e do conteúdo da publicação no seu conjunto, interpretando os vários artigos, números e partes e articulando-os num todo. Do ponto de vista prático, é acompanhado da recolha sistemática de dados concretos relativos a cada artigo, que permitirão posteriormente construir as bases de dados de termos analíticos para a revista.

Quais são, então, os dados que importa recolher? Em primeiro lugar, toda a informação identificativa do artigo: título, autor e localização na revista (número, secção e páginas). Em seguida, as referências feitas pelo articulista ao longo do texto: todas as menções a nomes, títulos (de livros, outros periódicos, etc.), instituições ou entidades colectivas, acontecimentos e locais geográficos. Noutro plano, são também identificados os termos empregues pelo autor para expor as suas ideias e conceitos, fundamentais como base de construção do ideário apresentado.

Esta informação é inserida na base de dados, agrupada por fichas, onde cada ficha representa um artigo distinto. Por outras palavras, a unidade elementar da base, que contém todo o universo referencial da revista em questão, é o artigo, aqui entendido não apenas como texto de opinião, mas também englobando notícias, poemas, contos, peças gráficas, publicidade, etc. Veja-se a figura da página seguinte, que descreve de forma esquemática os vários tipos de entrada numa ficha. Em itálico, encontram-se os dados que não são meramente textuais, ou seja, que requerem o contributo crítico do investigador.

Dado que, em quase todos os casos, o texto das fontes não se encontra disponível em formato digital, o que seria demasiado dispendioso e demorado obter, a automação de processos encontra-se limitada à organização e agilização da recolha de dados analíticos. A utilização de técnicas como o *text mining*, ou outras relacionadas com a computação semântica, é, assim, dificultada<sup>4</sup>. No entanto, e dado que se recorre a um conjunto de bases de dados integradas e uniformizadas, a pesquisa e a visualização de resultados de analíticos são inteiramente automatizadas.

<sup>4</sup> Esta via, caso existissem as fontes com as características necessárias, teria o potencial de produzir resultados ricos e proveitosos, desde que se respeitassem os princípios metodológicos relevantes.

**Figura 1: Composição de ficha**

Note-se que é feita uma distinção entre “conceito” e “assunto”: ao passo que o primeiro pressupõe uma maior profundidade conceptual<sup>5</sup>, o segundo pretende-se eminentemente descritivo e factual. Em ambos os casos, não basta atentar na referência textual explícita, ou seja, podem ser indexados conceitos e assuntos cujo sentido seja identificado pelo investigador, sem nunca serem referidos directamente no artigo.

A base é depois sujeita a uma série de revisões tanto sucessivas como exaustivas, procurando não apenas corrigir gralhas e erros, mas também uniformizar as entradas, de forma a garantir a sua coesão formal e a finalizar o processo prévio à sua publicação.

A par do trabalho do Seminário, é feita, por instituições em parceria, a digitalização das páginas da revista que, finalmente, acompanharão a base de dados. Durante esta

<sup>5</sup> O SLHI elaborou, como instrumento de trabalho, um *thesaurus* de conceitos, discutido e revisto periodicamente, que serve como referência para a escolha de termos.

fase, são feitas as ligações informáticas entre base e imagens digitalizadas, podendo optar-se pelo reconhecimento de caracteres e consequente possibilidade adicional de pesquisa livre<sup>6</sup>. O formato final da edição depende da revista e dos destinatários, perspectivando-se a sua disponibilização na Internet ou em suporte ROM.

Desta descrição facilmente se depreende o grau de detalhe que se procura atingir. A maioria dos esforços de indexação e classificação bibliográfica serve-se sobretudo de palavras-chave para facilitar o acesso à informação do leitor/investigador (deixando de parte a pesquisa livre que, como já vimos, representa uma estratégia de acesso à informação distinta). Nesta abordagem, para além dos dados formais bibliográficos (autor, título, local, data, idioma, etc.), são atribuídos descritores aos artigos que permitem buscas genericamente definíveis como buscas por assunto. Aqui cabem também temas, autores e obras, mas sempre a um nível muito simplificado e geral, apontando somente no sentido da relevância mais alta<sup>7</sup>. O seu único objectivo é descrever os principais assuntos tratados no texto, ignorando todas as demais referências e assumindo que o leitor, caso assim o deseje, a devido tempo se dedicará a estudá-las quando aceder directamente ao texto. Este princípio de simplificação, até certo ponto inevitável ao tratar de grandes acervos documentais, hierarquiza linearmente os assuntos, disponibilizando, na prática, apenas aqueles que ultrapassam determinado valor numa escala imaginária de pertinência. Dir-se-ia que apenas a extrema relevância permite a inclusão. Este esforço cumpre o tradicional e importante papel da organização como meio de gestão: classifica de forma sintética, eliminando toda a informação que possa ser secundária ou provocar ruído, contrastando em absoluto com o paradigma da pesquisa livre.

O SLHI procura uma abordagem algo distinta ao trabalho de indexação e classificação, sendo que se podem referir duas grandes diferenças teóricas em comparação com o princípio previamente referido no que toca à metodologia: exaustividade e interpretação crítica.

Apesar de ser sempre necessária alguma medida de síntese (para que, *in extremis*, a indexação não seja feita com o texto integral!), pretende-se que o investigador tenha à sua disposição todo o corpo de referências utilizado pelo articulista. Assim, e por exemplo, não são recolhidos apenas os autores tratados em maior profundidade, mas

<sup>6</sup> Como foi o caso da revista *Alma Nacional*. Cf. *Alma Nacional*, CD-ROM, Colecção de Revistas de Ideias e Cultura, n.º 1, Lisboa, Fundação Mário Soares, 2004.

<sup>7</sup> Veja-se, por exemplo, a SciELO. O mesmo se verifica na representação descritiva das bibliotecas nacionais.

sim todos, exaustivamente. O mesmo se aplica a todos os outros critérios de pesquisa, como títulos de obras, locais, etc. Isto permite saber precisamente quantas vezes e onde surge determinada referência, e para cada ocorrência permite ainda cruzá-la com as demais referências do mesmo artigo. Pode dizer-se que, neste aspecto, oferece ao investigador um mapa de conteúdos da publicação, antes ou sem que seja necessária a sua consulta directa. Esta forma de metadados define-se, juntamente com a representação descritiva clássica, como “informação estruturada que descreve, explica, localiza ou, de alguma forma, facilita o acesso, uso ou gestão de um recurso de informação”<sup>8</sup> – simplesmente, pretende que a rede referencial à disposição do leitor seja o mais completa possível e não sintética.

A isto junta-se a informação biobibliográfica complementar: todos os nomes são acompanhados das datas de nascimento e morte e das ocupações pelas quais foram conhecidos à época da publicação do texto, tanto quanto for possível apurá-las. São feitas pesquisas em todos os principais repositórios de dados biográficos, como catálogos de bibliotecas nacionais, enciclopédias e, caso seja necessário, dicionários especializados, para garantir que a informação não só está correcta, como é o mais abrangente possível. A Internet, para além de permitir aceder rapidamente a estes dados, representa também um recurso indispensável para chegar a referências mais obscuras e omitidas de todos os instrumentos clássicos<sup>9</sup>. Apesar de os nomes constituírem o objecto principal desta preocupação de documentação, ela estende-se também, ainda que em menor grau de aprofundamento, aos restantes tipos de entradas: datas e locais de congressos, cerimónias, publicações, etc. É sobre este trabalho que incide grande parte das revisões finais. É essencial, para o sucesso da utilização da base, garantir a sua homogeneidade e uniformidade, eliminando erros de recolha e mesmo de citação na própria revista, onde por vezes, por descuido ou por gralha, são feitas citações da mesma referência de forma díspar. Deve ainda dizer-se que o trabalho de revisão, considerando não apenas as incorrecções descritas mas também outros factores de dissonância (*e.g.*, uma grafia arcaica), dota o modelo de uma considerável vantagem sobre a pesquisa livre, dada a vulnerabilidade desta ao erro humano<sup>10</sup>.

<sup>8</sup> National Information Standards Organization, *Understanding Metadata*, NISO Press, 2004, p. 1.

<sup>9</sup> Mantém-se aqui algum equilíbrio entre relevância do nome e esforço necessário para recolher as informações pertinentes. Exceptuando alguns casos extremos, como nomes de familiares ou autores de uma obra apenas, pode dizer-se que a taxa de sucesso é muito elevada.

<sup>10</sup> Jeffrey Garrett, *op. cit.*

A outra particularidade do trabalho do SLHI prende-se com a intenção de ir além da simples disponibilização de informação: as bases de dados contêm termos não só de indexação, mas também de classificação dos conceitos utilizados. O conteúdo das revistas e seus artigos é, assim, sujeito a um olhar interpretativo crítico – daí a designação. O investigador não se limita a retirar algumas palavras e expressões do texto, *ipsis verbis*, como marcadores de indexação. Munido dos seus conhecimentos acerca dos assuntos tratados e do contexto histórico das revistas, dá um sentido focalizado aos conteúdos. Pretende-se que o seu contributo não seja meramente mecânico: as referências nas revistas não devem ser apenas identificadas e recolhidas, mas também interpretadas, articuladas e documentadas – alguns dos passos fundamentais de qualquer estudo historiográfico. Estes conceitos não coincidem, em definição, com os termos de indexação. Ao contrário destes últimos, é possível que a palavra que traduz o conceito nem sequer se encontre no texto – é responsabilidade do SLHI encontrar o conceito essencial que se encontra detrás da expressão textual de uma ideia.

Podemos, deste modo, resumir os principais objectivos do trabalho de edição de revistas em três pontos: permitir o acesso à informação; criar um aparato crítico que possibilite a sua análise detalhada; propor um estudo dos dados com base na experiência e nos instrumentos forjados. Se os princípios metodológicos apresentam, como acabámos de ver, acrescentos relativamente à indexação e classificação tradicionais ou mesmo à indexação e classificação electrónica emergentes, é inevitável que o horizonte dos objectivos também possa ser alargado. Aproveitando essa abertura, isto significa que, para além da acessibilidade, que se perspectiva em todos os casos, passa também a existir um instrumento de trabalho, não apenas de pesquisa. O resultado final da edição vai para além da criação de um meio de acesso à informação, concretiza-se na disponibilização de um instrumento de estudo<sup>11</sup>. A monografia ou o texto historiográfico que surge da mão do investigador após a finalização da base representa o culminar do processo, expressão do seu desenlace expectável e desejado: aproveitar o aparato crítico que se disponibiliza aos leitores para efectuar a sua primeira utilização especializada.

Pode-se, aliás, colocar inversamente a questão: a base de dados e a sua indexação e classificação são o trabalho necessário para, seguindo critérios metodológicos particulares e exaustivos, estudar determinada revista; uma vez que esse trabalho se

<sup>11</sup> Veja-se, como exemplo, António Nóvoa e Filomena Bandeira (coord.), *A Educação Portuguesa: Corpus documental (séculos XIX-XX)*, CD-ROM, Porto, Edições Asa, 2005. O CD reúne informação proveniente de diversas bases de dados bibliográficas relativas à área da educação, e a sua metodologia contém uma aproximação simplificada aos objectivos do SLHI.



encontre completo, e visto que se mantém válido e – julgamos – útil para futuros investigadores, será proveitosa a sua partilha.

Em jeito de comparação, podem ser destacadas algumas iniciativas de índole semelhante, apesar das diferenças metodológicas, como o trabalho levado a cabo para a revista *Colóquio/Letras*<sup>12</sup>, ou, a nível internacional, a edição em DVD da revista *Esprit*<sup>13</sup>. Deve notar-se que, nestes casos, é tratada apenas uma publicação, o que possibilita uma moldagem precisa de métodos e aparato crítico às suas características individuais. No entanto, o âmbito do SLHI é todo um universo de revistas, composto de constelações ideológicas coerentes, pelo que a sua estrutura metodológica deverá suportar variações particulares, sem condicionamento na sua aplicabilidade.

Se até aqui tem sido feita uma análise das sobreposições e especificidades do Seminário com os métodos de indexação e classificação geralmente utilizados, deve ser ainda mencionado um outro aspecto que, não sendo tão relevante nessa comparação, é fulcral para a caracterização em curso: o trabalho do SLHI é feito sobre revistas a partir de uma perspectiva da história das ideias, onde a noção de totalidade da revista é fundamental. Cada publicação encerra um conjunto de ideias, definidas e organizadas internamente de modos particulares. Identificar estes universos, articulá-los com os demais e situá-los no devido contexto histórico é um desígnio transversal do trabalho dos investigadores; ou seja, apesar de se tratar de revistas, dos seus impulsionadores e de todos os dados históricos que com eles se relacionam, as ideias que defendem e difundem estão sempre presentes como fio condutor e objecto final.

Retirando a componente analítica e historiográfica do trabalho, e cingindo-nos aos aspectos de indexação e classificação, é pertinente questionarmo-nos acerca das virtudes ou defeitos do modelo do Seminário relativamente aos restantes. Neste capítulo, pode dizer-se que a sua grande desvantagem é a considerável quantidade de trabalho – e portanto de tempo – que construir e finalizar uma base de dados deste tipo pressupõe. De forma a garantir a homogeneidade e correcção da base, são necessárias múltiplas revisões, técnicas e científicas, para além do trabalho volumoso que significa a própria recolha. Como seria de esperar, a extensão da revista em si, a quantidade de números, o tamanho dos mesmos, etc., determinam em grande parte o investimento que deve ser feito para a sua conclusão. Não obstante as eventuais complexidades particulares no ideário, o volume, e portanto o trabalho prático, de,

<sup>12</sup> Cf. <http://coloquio.gulbenkian.pt>.

<sup>13</sup> Cf. *La Collection intégrale d'Esprit, 1932-2006*, DVD-ROM, Revue Esprit, 2008.

por exemplo, *Orpheu* não é comparável ao da *Seara Nova*. Ainda assim, tratar uma qualquer revista com estes critérios, mesmo que tenham sido dela publicados apenas alguns números curtos, implica sempre um maior esforço do que o utilizado para a representação descritiva do mesmo título com fins de biblioteconomia.

Apesar de menos evidente, deve também ser mencionada a dificuldade do modelo em determinar hierarquias internas de relevância referencial. A unidade privilegiada de estudo do SLHI é, assumidamente, a revista na sua totalidade, ou seja, o conjunto de textos que a compõem. Assim, é fácil verificar quantitativamente a ocorrência de referências ao longo da publicação: a pesquisa por um nome ou por um conceito informa o leitor não só dos sítios onde se encontram, mas também do número total de entradas. No caso de *O Tempo e o Modo*<sup>14</sup>, o conceito “Fé” surge 33 vezes, ao passo que o conceito “Jogo” é utilizado apenas três vezes, indicando uma clara primazia do primeiro sobre o segundo em termos de importância no ideário da publicação. Aferir a mesma informação para apenas um artigo é, no entanto, mais problemático, já que, a partir da base, não existem dados que permitam fazê-lo imediatamente, sem a leitura do próprio texto. Os descritores biblioteconómicos também não resolvem esta insuficiência, mas, como geralmente procuram a síntese qualitativa de conteúdos, a sua natureza conduz o leitor aos temas mais importantes do texto, apesar de o fazer com o sacrifício de todos os outros. Neste aspecto e à primeira vista, o modelo do SLHI não parece ser tão assinalável a lidar com artigos individuais como com o conjunto da publicação. No entanto, esta dificuldade é facilmente ultrapassada fazendo o cruzamento com os restantes dados contidos na ficha – contextualizando o resultado automático da busca, depressa se determina o peso relativo no artigo. Mais uma vez, o principal motivo para isto suceder é o objecto globalmente mais importante de trabalho de um e outro modelos: texto/artigo/livro vs. revista/totalidade de uma publicação, e as consequências daí resultantes.

Nenhuma destas aparentes limitações resulta de falhas metodológicas estruturais. Decorrem das opções e dos objectivos do modelo analítico escolhido e são, nesse quadro, inevitáveis. Como, no âmbito do trabalho do Seminário, a indexação e a classificação não representam um fim, mas antes um meio – ao contrário do trabalho de representação descritiva das bibliotecas ou das iniciativas de edição electrónica de periódicos –, a sua operacionalização acarreta consequências particulares, tanto posi-

<sup>14</sup> *O Tempo e o Modo – 1.ª série*, DVD-ROM, Colecção de Revistas de Ideias e Cultura, n.º 2, Lisboa, Fundação Mário Soares, 2006.

tivas como negativas. Por este motivo, a comparação nunca pode ser feita de forma totalmente directa: os instrumentos de indexação e classificação partilham linguagens semelhantes e coincidem em alguns pontos, mas os objectivos assumidos num e noutro caso determinam, de facto, aplicações práticas distintas.

A actualidade da edição electrónica de periódicos caminha, a par do desenvolvimento e difusão de meios informáticos cada vez mais eficazes, na direcção da acessibilidade alargada à informação. O SLHI propõe uma utilização especializada e aprofundada desses instrumentos, para a construção de uma nova metodologia de edição electrónica *analítica*. Neste sentido, apresenta uma abordagem própria, moldada pelas suas raízes na História e na Filosofia, aos princípios de uma das mais interessantes fronteiras tecnológicas da actualidade, a da computação semântica, aplicada à edição electrónica. Edição e estudo deixam de constituir etapas inteiramente distintas no processo de tratamento de uma revista – histórica ou não, apesar de, no caso do Seminário, esta característica ser fundamental –, para passarem a formar, desde o primeiro momento e de forma integrada, um só objectivo simbiótico.